ELSEVIER
Linear Algebra and its Applications 279 (1998) 39–50
LINEAR ALGEBRA AND ITS APPLICATIONS

On Schur D-stable matrices

R. Fleming *, G. Grossman, T. Lenker, S. Narayan [1], S.-C. Ong

Department of Mathematics, Central Michigan University, Mount Pleasant, MI 48859, USA

Received 1 November 1996; accepted 23 November 1997

Submitted by D.H. Carlson

Abstract

It is shown that vertex stability implies *Schur D-stability* for real 2×2 matrices and real $n \times n$ tridiagonal matrices. Additional results describing the class of $n \times n$ complex *Schur D-stable* matrices are given.

AMS classification: 15A18; 47A10

Keywords: Schur stable; Schur D-stable; Vertex stable

1. Introduction

A square complex matrix A is said to be *Schur stable* if the spectrum of A is contained in the open unit disk in the complex plane. This property plays an important role in the stability theory for discrete time dynamical systems. In this paper we explore a related concept of stability of matrices called *Schur D-stability* which occurs naturally in the area of control systems analysis and design [1,9]. For additional information on matrix stability consult [2,3,5].

Given a square complex matrix $A = [a_{ij}]$, $|A|$ denotes the matrix $[|a_{ij}|]$. For a real matrix A, $A \geqslant 0$ (resp. $A > 0$) means that all entries of A are nonnegative (resp. positive). For two real matrices A and B, we write $A \leqslant B$ or $B \geqslant A$ for

* Corresponding author. Tel.: +1 517 7743238; fax: +1 517 7742414; e-mail: richard.fleming@cmich.edu.

[1] Author gratefully acknowledges the support given by Central Michigan University during his sabbatical leave.

0024-3795/98/$19.00
PII: S0024-3795(98)10058-7

$B - A \geqslant 0$. The set of all eigenvalues of A, denoted by $\sigma(A)$, is called the spectrum of A. The spectral radius of A is denoted by $\rho(A)$ $(= \max\{|\lambda|: \lambda \in \sigma(A)\})$. The operator norm of A is denoted by $||A||$, which is $\sqrt{\rho(A^*A)}$, where A^* is the adjoint (conjugate transpose) of A. The algebra of $n \times n$ complex (real) matrices with operator norm will be denoted by $\mathscr{M}_n(\mathbb{C})$ (resp. $\mathscr{M}_n(\mathbb{R})$). The identity matrix of size appropriate for the context will be denoted by I.

Definition 1.1. A real or complex square matrix A is said to be
(a) *Schur stable* if $\rho(A) < 1$.
(b) *Schur D-stable* if $\rho(AD) < 1$ for every real diagonal D with $|D| \leqslant I$.
(c) *vertex stable* if $\rho(AD) < 1$ for all real diagonal D with $|D| = I$.

We will denote the set of all $n \times n$ real (resp. complex) Schur D-stable matrices by $\mathscr{S}\mathscr{D}_n(\mathbb{R})$ (resp. $\mathscr{S}\mathscr{D}_n(\mathbb{C})$).

2. Real 2 × 2 Schur D-stable matrices

In this section we give a characterization of the set $\mathscr{S}\mathscr{D}_2(\mathbb{R})$. First we give a characterization of Schur stable 2×2 matrices.

Proposition 2.1. *Let*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathscr{M}_2(\mathbb{R}).$$

Then A is Schur stable if and only if
(i) $|ad - bc| < 1$, *and*
(ii) $|a + d| < 1 + (ad - bc)$.

Proof. Suppose that A is Schur stable. The eigenvalues of A are

$$\lambda_\pm = \frac{(a+d) \pm \sqrt{(a+d)^2 - 4(ad-bc)}}{2}, \tag{2.1}$$

which must satisfy $|\lambda_\pm| < 1$. Since $\det(A) = (\lambda_+)(\lambda_-)$, (i) follows.

If $(a+d)^2 - 4(ad-bc) \geqslant 0$, we show that condition (ii) of the proposition follows from $|\lambda_\pm| < 1$. Let us further assume that $a + d < 0$ and leave the other case $a + d \geqslant 0$ to the reader. Then we get

$$-(a+d) + \sqrt{(a+d)^2 - 4(ad-bc)} < 2.$$

This implies

$$(a+d)^2 - 4(ad-bc) < [2 + (a+d)]^2 = 4 + 4(a+d) + (a+d)^2$$

giving $|a+d| = -(a+d) < 1 + (ad-bc)$ as desired.

If $(a+d)^2 - 4(ad-bc) < 0$ in (2.1) then $(a-d)^2 < -4bc$. Thus

$$\begin{aligned} 1 + (ad-bc) &> 1 + ad + \frac{(a-d)^2}{4} = \frac{4+(a+d)^2}{4} \\ &\geqslant \frac{2\cdot 2|a+d|}{4} = |a+d|. \end{aligned}$$

Condition (ii) holds in all cases, and one implication is proved.

Now assume that A satisfies the conditions in the statement of the proposition. To see that the eigenvalues $\lambda_\pm$ are in the open unit disk, we consider two cases again. Suppose $(a+d)^2 - 4(ad-bc) \geqslant 0$, and $a+d<0$. By (ii) we have $-4(a+d) < 4 + 4(ad-bc)$, giving

$$(a+d)^2 - 4(ad-bc) < 4 + 4(a+d) + (a+d)^2 = [2+(a+d)]^2. \tag{2.2}$$

It follows easily from (i) and (ii) that $|a+d|<2$. Using this together with Eqs. (2.1) and (2.2), we obtain the inequality

$$\begin{aligned} -1 = \frac{(a+d) - \sqrt{(a+d+2)^2}}{2} &< \lambda_- < \lambda_+ \\ < \frac{(a+d) + \sqrt{(a+d+2)^2}}{2} &= 1 + (a+d) < 1 \end{aligned}$$

from which $|\lambda_\pm| < 1$ is clear.

In the case of $a+d \geqslant 0$, a similar argument using $a+d<2$ yields $|\lambda_\pm| < 1$. If $(a+d)^2 - 4(ad-bc) < 0$, then $|\lambda_+|^2 = (ad-bc) < 1$, by (i). This completes the proof. □

The following combines results found in [1,9].

Theorem 2.2. *Let*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathcal{M}_2(\mathbb{R}).$$

Then $A \in \mathcal{SD}_2(\mathbb{R})$ *if and only if*

(i) $|ad-bc| < 1$,

(ii) $|a+d| < 1 + (ad-bc)$, *and*

(iii) $|a-d| < 1 - (ad-bc)$.

That is, A is Schur D-stable if and only if A is vertex stable.

Proof. Assume that A is Schur D-stable. Then A is vertex stable, and conditions (i) and (ii) follow from Proposition 2.1 since A is Schur stable. The matrix

$$B = \begin{bmatrix} a & -b \\ c & -d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

is also Schur stable; so that (iii) follows from Proposition 2.1 again.

Suppose now that A satisfies conditions (i)–(iii). Let

$$D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$$

with entries d_1, d_2 in $[-1, 1]$. Under the assumptions that $d_1 \neq 0$ and $|d_2| \leqslant |d_1|$, we see that

$$\left|\frac{1}{d_1}D\right| = \begin{bmatrix} 1 & 0 \\ 0 & |d_2/d_1| \end{bmatrix} \leqslant I,$$

and

$$\rho\left[A\left(\frac{1}{d_1}D\right)\right] = \frac{1}{|d_1|}\rho(AD) \geqslant \rho(AD).$$

Hence it is enough to consider

$$K = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix} \quad \text{and} \quad K = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$$

with $|k| \leqslant 1$. We will consider the first case and leave the other case for the reader. We must show that

$$AK = \begin{bmatrix} a & bk \\ c & dk \end{bmatrix}$$

satisfies the conditions in Proposition 2.1. The first condition is clear. The second is $|a + kd| < 1 + k(ad - bc)$. Unless $a + kd = 0$ for some k with $|k| < 1$, both sides are linear functions of k. In the other case $|a + kd|$ is piecewise linear, and concave up. Since $|a + kd| < 1 + k(ad - bc)$ holds at $k = \pm 1$ by (ii) and (iii), this inequality must hold throughout the interval. Hence, A is vertex stable implies A is Schur D-stable. Note that conditions (i)–(iii) imply A is vertex stable by Proposition 2.1. □

3. Real 3 × 3 Schur stable matrices

Let $A = [a_{ij}] \in \mathscr{M}_3(\mathbb{R})$. We begin this section by stating necessary and sufficient conditions that A be Schur stable. Our second result characterizes all 3×3 matrices which are vertex stable.

The 2×2 principal minors of A will be denoted by $m_{11} = a_{22}a_{33} - a_{23}a_{32}$, $m_{22} = a_{11}a_{33} - a_{13}a_{31}$ and $m_{33} = a_{11}a_{22} - a_{12}a_{21}$. The trace and the determinant of A are denoted by τ and δ, respectively. The characteristic polynomial of A becomes $p(\lambda) = -\lambda^3 + \tau\lambda^2 - \mu\lambda + \delta$, where $\mu = m_{11} + m_{22} + m_{33}$.

Theorem 3.1. *Let $A \in \mathscr{M}_3(\mathbb{R})$. Then, using the above notation, A is Schur stable if and only if*

(i) $|\delta| < 1$,

(ii) $|\tau + \delta| < 1 + \mu$,

(iii) $|\tau\delta - \mu| < 1 - \delta^2$.

Furthermore (i)–(iii) *imply*

(iv) $(1 - \delta^2 + \tau^2 - \mu^2)(1 - \delta^2 - \delta\tau + \mu) > 2(\delta\mu - \tau)^2$.

Proof. Assume A is Schur stable.

(i) If x, y, z are the eigenvalues of A, then $|x|, |y|, |z| < 1$. Therefore $|\delta| < 1$.

(ii) The real eigenvalues of A lie strictly between -1 and 1. The cubic polynomial $p(\lambda)$ satisfies $\lim_{\lambda\to-\infty} p(\lambda) = +\infty$ and $\lim_{\lambda\to\infty} p(\lambda) = \infty$. It follows that $p(-1) = 1 + \tau + \mu + \delta > 0$ and $p(1) = -1 + \tau - \mu + \delta < 0$. Hence $|\tau + \delta| < 1 + \mu$.

(iii) *Case* 1. $\tau\delta - \mu \geqslant 0$ and $\delta > 0$. Then using $\tau + \delta < 1 + \mu$ from (ii), we get $\tau\delta + \delta^2 < \delta + \delta\mu$. Hence $0 \leqslant \tau\delta - \mu < \delta + \delta\mu - \mu - \delta^2 = (1 - \delta)(-\mu + \delta)$, so $0 \leqslant \tau\delta - \mu < 1 - \delta^2$ since $1 - \delta > 0$ from (i) and $\mu > -1$ from (ii).

Case 2. $\tau\delta - \mu \geqslant 0$ and $\delta < 0$. Then using $\tau + \delta > -1 - \mu$ from (ii), we get $\tau\delta + \delta^2 < -\delta - \delta\mu$. Therefore

$$0 \leqslant \tau\delta - \mu < -\delta - \delta\mu - \mu - \delta^2 = (1 + \delta)(-\mu - \delta) < 1 - \delta^2.$$

Case 3. $\tau\delta - \mu \geqslant 0$ and $\delta = 0 \Rightarrow \mu \leqslant 0$. From (ii) we have $0 < 1 + \mu \Rightarrow -1 < \mu \leqslant 0$ so that $|\mu| < 1$.

Case 4. $\tau\delta - \mu < 0$. Assume that x, y, z are also real. Then A is similar to the matrix

$$\begin{bmatrix} x & p & 0 \\ 0 & y & q \\ 0 & 0 & z \end{bmatrix}$$

in Jordan canonical form, where p and q are 0 or 1. By the invariance of the characteristic polynomial under similarity, we have $\delta = xyz$, $\tau = x + y + z$, and $\mu = xy + xz + yz$. We wish to show that $\delta^2 - 1 < \tau\delta - \mu < 0$ or $-\tau\delta + \mu < 1 - \delta^2$. In terms of the above Jordan matrix we must show $-xyz(x + y + z) + xy + xz + yz < 1 - x^2y^2z^2$. Observe that $f(x,y,z) = 1 - x^2y^2z^2 + x^2yz + xy^2z + xyz^2 - xy - xz - yz = (1 - xy)(1 - xz)(1 - yz)$ which is positive for all real x, y, z with $|x|, |y|, |z| < 1$. Therefore the desired inequality holds when the eigenvalues x, y, z are real. Since we are considering 3×3 real matrices, we may assume that two of the eigenvalues are complex conjugates of each other and one eigenvalue is real. If x is real, y and z are complex conjugates of each other, then $f(x,y,z) = |1 - xz|^2(1 - |z|^2)$ which is positive for $|x|, |y|, |z| < 1$. In all the cases we have $|\tau\delta - \mu| < 1 - \delta^2$.

(iv) From (iii) we have $1 - \delta^2 + \delta\tau - \mu > 0$, and squaring (ii) we get $1 - \delta^2 - 2(\tau\delta - \mu) > \tau^2 - \mu^2$. Thus,

$$(1 - \delta^2 + \delta\tau - \mu)[1 - \delta^2 - 2(\delta\tau - \mu)] - (1 - \delta^2 + \delta\tau - \mu)(\tau^2 - \mu^2) > 0.$$

After expanding the left side, using the fact that $(\delta\tau - \mu)^2 = (\delta\mu - \tau)^2 - (1 - \delta^2)(\tau^2 - \mu^2)$, and simplifying, we have,

$$\begin{aligned}&(1 - \delta^2)^2 - (1 - \delta^2)(\delta\tau - \mu) + (1 - \delta^2)(\tau^2 - \mu^2)\\ &\qquad - (\delta\tau - \mu)(\tau^2 - \mu^2) - 2(\delta\mu - \tau)^2 > 0,\end{aligned}$$

which is readily seen to be equivalent to (iv) after a factorization.

To prove that (i)–(iii) are sufficient for Schur stability of A we apply the Schur–Cohn conditions ([8], p. 152) to the characteristic polynomial for A. Since the conditions (i)–(iii) imply (iv), we note the following:

(A) $\Delta_1 = \det\left(\begin{bmatrix} \delta & -1 \\ -1 & \delta \end{bmatrix}\right) < 0 \iff |\delta| < 1$ which is (i),

(B) $\Delta_2 = \det\left(\begin{bmatrix} \delta & 0 & -1 & \tau \\ -\mu & \delta & 0 & -1 \\ -1 & 0 & \delta & -\mu \\ \tau & -1 & 0 & \delta \end{bmatrix}\right) > 0 \iff |\tau\delta - \mu| < 1 - \delta^2$

which is (iii),

(C) $\Delta_3 = \det\left(\begin{bmatrix} \delta & 0 & 0 & -1 & \tau & -\mu \\ -\mu & \delta & 0 & 0 & -1 & \tau \\ \tau & -\mu & \delta & 0 & 0 & -1 \\ -1 & 0 & 0 & \delta & -\mu & \tau \\ \tau & -1 & 0 & 0 & \delta & -\mu \\ -\mu & \tau & -1 & 0 & 0 & \delta \end{bmatrix}\right) < 0 \iff$ (iv).

To evaluate the determinant Δ_3, we take the transpose of the matrix, label the new matrix as a block matrix of 3×3 blocks

$$\begin{bmatrix} X & Y \\ Z & W \end{bmatrix},$$

and notice that Z and W commute. A theorem of Schur (see [4], p. 39) says that the determinant Δ_3 of the block matrix is the same as the determinant of the matrix $XW - YZ$. After simplification, we obtain

$$\Delta_3 = [(\delta^2 - 1 + \mu^2 - \tau^2)(\delta^2 - 1 + \delta\tau - \mu) - 2(\tau - \delta\mu)^2](\delta^2 - 1 - \delta\tau + \mu)$$

which is negative exactly when (iv) is valid, under our other assumptions. Therefore the eigenvalues of A are inside the unit circle, and the proof is complete. $\square$

Theorem 3.2. *Let $A \in \mathcal{M}_3(\mathbb{R})$. Then A is vertex stable if and only if the following conditions are valid for A:*

(i) $|\det(A)| < 1$,
(ii) $|a_{11} + a_{22} + a_{33} + \det(A)| < 1 + m_{11} + m_{22} + m_{33}$,
(iii) $|a_{11} - a_{22} - a_{33} + \det(A)| < 1 + m_{11} - m_{22} - m_{33}$,
(iv) $|a_{11} + a_{22} - a_{33} - \det(A)| < 1 - m_{11} - m_{22} + m_{33}$,
(v) $|a_{11} - a_{22} + a_{33} - \det(A)| < 1 - m_{11} + m_{22} - m_{33}$,
(vi) $|\det(A)(a_{11} + a_{22} + a_{33}) - (m_{11} + m_{22} + m_{33})| < 1 - [\det(A)]^2$,
(vii) $|\det(A)(-a_{11} - a_{22} + a_{33}) + (m_{11} + m_{22} - m_{33})| < 1 - [\det(A)]^2$,
(viii) $|\det(A)(-a_{11} + a_{22} - a_{33}) + (m_{11} - m_{22} + m_{33})| < 1 - [\det(A)]^2$,
(ix) $|\det(A)(a_{11} - a_{22} - a_{33}) + (-m_{11} + m_{22} + m_{33})| < 1 - [\det(A)]^2$.

where m_{jj}, $j = 1, 2, 3$, are principal minors of A as defined in the paragraph preceding the statement of Theorem 3.1.

Proof. Apply Theorem 3.1 to the matrices AD with D a real diagonal matrix satisfying $|D| = I$. Note the conditions (i) and (iii) of Theorem 3.1 applied to AD yield the same results for $A(-D)$. $\square$

Using Theorem 3.1, a result about the Schur D-stability of A can also be stated as follows: A 3×3 real matrix A is Schur D-stable if and only if for each diagonal matrix D with $|D| \leqslant I$ the matrix $B = AD$ satisfies the conditions (i)–(iii) in Theorem 3.1.

4. Tridiagonal Schur D-stable matrices

In this section we prove that vertex stability implies Schur D-stability for certain block triangular matrices and tridiagonal matrices.

Lemma 4.1. *Let $A \in \mathcal{M}_n(\mathbb{C})$ be block triangular with square blocks down its diagonal. Then*

(i) *A is vertex stable if and only if each of the blocks on the diagonal is vertex stable.*

(ii) *A is Schur D-stable if and only if each of the blocks on the diagonal is Schur D-stable.*

The proof of this lemma is straightforward. The main theorem of this section is the following,

Theorem 4.2. *Let $A \in \mathcal{M}_n(\mathbb{R})$ be a tridiagonal matrix. Then A is Schur D-stable if and only if A is vertex stable.*

Proof. By applying Lemma 4.1 (repeatedly, if necessary) we need only consider the case with the super- and sub-diagonal entries of A all nonzero. Starting with $d_1 = 1$ select $d_{j+1} = \pm 1$ inductively so that $a_{j+1,j}d_j a_{j,j+1}d_{j+1} > 0$. Next choose $p_1 = 1$ and inductively set $p_{j+1} = p_j\sqrt{a_{j+1,j}d_j/(a_{j,j+1}d_{j+1})} > 0$. If $P = \mathrm{diag}(p_1, p_2, \dots, p_n)$ and $D = \mathrm{diag}(d_1, d_2, \dots, d_n)$, then a straightforward calculation shows that $P^{-1}ADP$ is symmetric. Thus,

$$||P^{-1}ADP|| = \rho(P^{-1}ADP) = \rho(AD) < 1$$

since we assume A is vertex stable. If $|K| \leqslant I$, then

$$\begin{aligned}\rho(AK) &= \rho(P^{-1}ADD^{-1}KP) = \rho(P^{-1}ADPD^{-1}K)\\ &\leqslant ||P^{-1}ADP||\,||D^{-1}K|| < 1.\end{aligned}$$

which implies A is Schur D-stable. The converse is clear. □

Proposition 4.3. *If $A \in \mathcal{M}_n(\mathbb{R})$ is block triangular with square tridiagonal blocks down its diagonal then A is Schur D-stable if and only if A is vertex stable.*

Proof. Apply Lemma 4.1(i) followed by Theorem 4.2 and Lemma 4.1(ii). □

5. General results on Schur D-stable matrices

In this section we prove some general results concerning the class $\mathcal{SD}_n(\mathbb{C})$. We also give some examples and general remarks.

Theorem 5.1. *If A is in $\mathcal{SD}_n(\mathbb{C})$ (or $\mathcal{SD}_n(\mathbb{R})$), then there exists a $\delta > 0$ such that $\rho(AD) \leqslant 1 - \delta$ for every diagonal D with $|D| \leqslant I$.*

Proof. Since the set of all diagonal matrices D with $|D| \leqslant I$ is compact, and the matrix multiplication operation is continuous in the norm topology, the set $\mathcal{K}_A = \{AD\colon |D| \leqslant I\}$ is compact. As the spectral radius $\rho : X \mapsto \rho(X)$ is a continuous map from $\mathcal{M}_n(\mathbb{C})$ (with operator norm topology) to the reals, it maps $\mathcal{K}_A$ to a compact subset of the reals. Thus $\rho(\mathcal{K}_A)$ has a maximum, $\rho(AD_0)$ for some D_0 with $|D_0| \leqslant I$. But since is Schur D-stable, $\rho(AD_0) < 1$. Choose $\delta = 1 - \rho(AD_0)$. Then we have the desired conclusion. □

Theorem 5.2. *$\mathscr{SD}_n(\mathbb{C})$ is open in $\mathscr{M}_n(\mathbb{C})$.*

Proof. Let $A \in \mathscr{SD}_n(\mathbb{C})$. Then $\mathscr{K}_A = \{AD: |D| \leqslant I\}$ is a compact subset of $\rho^{-1}((-\infty, 1))$. Since ρ is continuous, $\rho^{-1}((-\infty, 1))$ is open. Thus there exists a $\delta > 0$ such that if $\|X - B\| < \delta$ for some $B \in \mathscr{K}_A$, then we have $X \in \rho^{-1}((-\infty, 1))$. Now let X be such that $\|X - A\| < \delta$. Then for each D with $|D| \leqslant I$, we have $\|XD - AD\| \leqslant \|X - A\| < \delta$. Thus $XD \in \rho^{-1}((-\infty, 1))$, or $\rho(XD) < 1$; i.e., X is Schur D-stable. □

Theorem 5.3. *Let $\mathscr{G}$ be the subset of $\mathscr{M}_n(\mathbb{C})$ defined by*

$$\mathscr{G} = \{VP\colon\ P \text{ a permutation matrix and } V \text{ a diagonal unitary matrix}\}.$$

Let $A \in \mathscr{M}_n(\mathbb{C})$ be a triangular matrix with the modulus of each diagonal entry strictly bounded by 1. *Then there exists a $\delta > 0$ such that T^*AT is Schur D-stable for all matrices T satisfying the condition that* $\operatorname{dist}(T, \mathscr{G}) < \delta$.

We note that $\mathscr{G}$ is in fact a group, though we will not be using this.

Proof. Define $F(X) = X^*AX$ for all $X \in \mathscr{M}_n(\mathbb{C})$. Then F is continuous, and hence $\mathscr{V} = F^{-1}(\mathscr{SD}_n(\mathbb{C}))$ is open. It is easy to see that $\mathscr{G}$ is compact and is contained in $\mathscr{V}$. Thus there exists a δ such that for all T with $\operatorname{dist}(T, \mathscr{G}) < \delta$, we have $T \in \mathscr{V}$. This is equivalent to $T^*AT \in \mathscr{SD}_n(\mathbb{C})$ for all T with $\operatorname{dist}(T, \mathscr{V}) < \delta$, as desired. □

Note that the set of Schur D-stable matrices that are obtained as described in this theorem is quite large. We do not know if all Schur D-stable matrices arise in this way.

Theorem 5.4. *A matrix $A \in \mathscr{M}_n(\mathbb{C})$ is in $\mathscr{SD}_n(\mathbb{C})$ if and only if for every k, $1 \leqslant k \leqslant n$, every k-principal submatrix of A is in $\mathscr{SD}_k(\mathbb{C})$.*

Proof. Sufficiency is clear. To prove the necessity observe that for a $k \times k$ principal submatrix B of A, there exists a permutation matrix P such that

$$P^{t}AP = \begin{bmatrix} B & * \\ * & * \end{bmatrix}.$$

Let D be a $k \times k$ diagonal matrix with $|D| \leqslant I$. Let

$$\tilde{D} = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

be the $n \times n$ matrix obtained from D by adjoining zero rows and columns. Then $|\tilde{D}| \leqslant I$, and we have

$$\sigma(BD) \subseteq \sigma\left(\begin{bmatrix} B & * \\ * & * \end{bmatrix}\tilde{D}\right) = \sigma(P^{t}AP\tilde{D}) = \sigma(AP\tilde{D}P^{t}).$$

Since $P\tilde{D}P^{t}\ A$ is diagonal with $|P\tilde{D}P^{t}| \leqslant I$, and is Schur D-stable, it follows that $\sigma(BD)$ is contained in the open unit disk. Since D is an arbitrary $k \times k$ diagonal matrix with $|D| \leqslant I$, we conclude that $B \in \mathscr{S}\mathscr{D}_k(\mathbb{C})$. □

Lemma 5.5. *Let $A, B \in \mathscr{M}_n(\mathbb{C})$. Then*
(i) *If $|A| \leqslant B$, then $\rho(A) \leqslant \rho(|A|) \leqslant \rho(B)$.*
(ii) *If $0 \leqslant A \leqslant B$, then $\rho(A) \leqslant \rho(B)$.*
(iii) *If $A \geqslant 0$, then $A \in \mathscr{S}\mathscr{D}_n(\mathbb{C})$ if and only if A is Schur stable.*

Proof. The proofs of (i) and (ii) are found on p. 491 of [6]. For (iii), let $A \geqslant 0$ and suppose K is diagonal with $|K| \leqslant I$. Then

$$\rho(AK) \leqslant \rho(|AK|) = \rho(A|K|) \leqslant \rho(A),$$

by (i) and (ii) which proves (iii). □

Theorem 5.6. *Let $A \in \mathscr{M}_n(\mathbb{C})$. Suppose that*

$$P^{t}AP = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

for some permutation matrix P where $A_{11} \geqslant 0$ and $A_{22} \geqslant 0$ are square matrices and $A_{12} \leqslant 0$ and $A_{21} \leqslant 0$. Then $A \in \mathscr{S}\mathscr{D}_n(\mathbb{C})$ if and only if A is Schur stable.

Proof. Assume A_{11} is $j \times j$ and define $D = \operatorname{diag}(1, \ldots, 1, -1, \ldots, -1)$ with exactly j elements equal to 1. Then $B = DP^{t}APD = |P^{t}AP|$. If K is a diagonal matrix with $|K| \leqslant I$, then $\rho(AK) = \rho(DP^{t}AKPD) = \rho(DP^{t}APDDP^{t}KPD) = \rho(BL)$, where $L = DP^{t}KPD$ is a diagonal matrix which satisfies $|L| \leqslant I$. Since $|BL| \leqslant B$ we have $\rho(A) = \rho(BL) \leqslant \rho(B)$ by Lemma 5.5(i). Finally, $\rho(B) = \rho(DP^{t}APD) = \rho(A)$ which implies that $A \in \mathscr{S}\mathscr{D}_n(\mathbb{C})$ if A is Schur stable. The converse is immediate. □

Proposition 5.7. *$\mathscr{S}\mathscr{D}_n(\mathbb{C})$ is path-wise connected.*

Proof. If A and B are in $\mathscr{S}\mathscr{D}_n(\mathbb{C})$, A and B can be joined through the zero matrix since $tA + (1-t)0$ is in $\mathscr{S}\mathscr{D}_n(\mathbb{C})$ for $0 \leqslant t \leqslant 1$. □

Remark 5.8. (i) $\mathscr{S}\mathscr{D}_n(\mathbb{C})$ contains all matrices A such that $\|P^{-1}AP\| < 1$, for some invertible diagonal matrix P. In particular $\mathscr{S}\mathscr{D}_n(\mathbb{C})$ contains the open unit ball of $\mathscr{M}_n(\mathbb{C})$. This is because, for every diagonal K with $|K| \leqslant I$, we have $\rho(AK) = \rho(APKP^{-1}) = \rho(P^{-1}APK) \leqslant \|P^{-1}AP\| \cdot \|K\| < 1$.

(ii) If $n > 1$, $\mathcal{SD}_n(\mathbb{C})$ is unbounded. Since, if a_{12} is the only nonzero entry of A, $\|A\| = |a_{12}|$, and $\rho(AD) = 0$ for every diagonal D.

(iii) It is immediate from (i) that a matrix A with $\rho(A) = \|A\|$ is in $\mathcal{SD}_n(\mathbb{C})$ if and only if $\|A\| < 1$. Since normal matrices satisfy $\rho(A) = \|A\|$ we can therefore conclude that there are no normal matrices outside the unit ball that are Schur D-stable.

(iv) That $\mathcal{SD}_n(\mathbb{C})$ is not convex can be seen by the following example:

$$A = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}$$

are in $\mathcal{SD}_2(\mathbb{C})$, but

$$\frac{1}{2}A + \frac{1}{2}B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

has eigenvalues ± 1, and therefore is not in $\mathcal{SD}_2(\mathbb{C})$.

(v)

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

is in $\mathcal{SD}_2(\mathbb{C})$ but $A^{\mathrm{t}}A$ is not. In particular $\mathcal{SD}_n(\mathbb{C})$ is not closed under multiplication.

(vi) $\mathcal{SD}_n(\mathbb{C})$ is not closed under (orthogonal, or unitary) similarity. For example,

$$A = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

Then A is in $\mathcal{SD}_2(\mathbb{C})$, but

$$B^{\mathrm{t}}AB = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix},$$

which is not in $\mathcal{SD}_2(\mathbb{C})$ by Theorem 5.4.

6. Schur D-Stability and matrix norms

The following theorem is taken from [7]. See also p. 310 of [6].

Theorem 6.1. *Let $\mathcal{N}(.)$ be a vector norm on $\mathbb{C}^n$ and let $\tilde{\mathcal{N}}(.)$ be the induced operator norm on $\mathcal{M}_n(\mathbb{C})$. Then the following statements are equivalent.*

(a) *$\mathcal{N}(x) \leqslant \mathcal{N}(y)$ for all x, y in $\mathbb{C}^n$ with $|x| \leqslant |y|$ (such an $\mathcal{N}(.)$ is also known as an monotone norm).*

(b) $\mathcal{N}(x) = \mathcal{N}(|x|)$ *for all* x *in* $\mathbb{C}^n$ *(such an* $\mathcal{N}(.)$ *is also known as an absolute norm).*

(c) $\tilde{\mathcal{N}}(D) = \rho(D)$ *for* $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$.

(d) $\tilde{\mathcal{N}}(D) \leqslant 1$ *for* D *diagonal with* $|D| \leqslant I$.

Proposition 6.2. *Let* $A \in \mathcal{M}_n(\mathbb{C})$. *Each of the following conditions implies* A *is Schur D-stable.*

(a) *There exists a matrix norm* $\tilde{\mathcal{N}}(.)$ *such that* $\tilde{\mathcal{N}}(A) < 1$ *and* $\tilde{\mathcal{N}}(D) \leqslant 1$ *whenever* $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$ *with* $|d_j| \leqslant 1, j = 1, 2, \ldots, n$.

(b) *There exists a monotone norm* $\mathcal{N}(.)$ *on* $\mathbb{C}^n$ *such that* $\tilde{\mathcal{N}}(A) < 1$.

(c) *There exists a monotone norm* $\mathcal{N}(.)$ *on* $\mathbb{C}^n$ *such that* $\tilde{\mathcal{N}}(AD) < 1$ *for* $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$ *with* $|d_j| = 1$, *for* $j = 1, 2, \ldots, n$.

Proof. (a) $\rho(AD) \leqslant \tilde{\mathcal{N}}(AD) \leqslant \tilde{\mathcal{N}}(A) \cdot \tilde{\mathcal{N}}(D) < 1$. (c) is a special case of (b) which is a special case of (a). □

Acknowledgements

We thank Professor Bryan Cain for introducing the problem to one of the authors. We thank the referee for many helpful suggestions.

References

[1] A. Bhaya, E. Kaszkurewicz, On discrete-time diagonal and D-stability, Linear Alg. Appl. 187 (1993) 87–104.
[2] G.W. Cross, Three types of matrix stability, Linear Alg. Appl. 20 (1978) 253–263.
[3] Gene H. Golub, Charles Van Loan, Matrix Computations, John Hopkins University Press, Baltimore, MD, 1983.
[4] P. Halmos, A Hilbert Space Problem Book, Springer, New York, 1982.
[5] D. Hershkowitz, Recent directions in matrix stability, Linear Alg. Appl. 171 (1992) 161–186.
[6] Roger A. Horn, Charles R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1993.
[7] C.R. Johnson, P. Nylen, Monotonicity properties of norms, Linear Alg. Appl. 148 (1991) 43–58.
[8] M. Marden, Geometry of the Zeros of a Polynomial in a Complex Variable, Math Surveys III, American Mathematical Society, Providence, RI, 1949.
[9] W.L. Mills, C.T. Mullis, R.A. Roberts, Digital filter realizations without overflow oscillations, IEEE Trans. Acoustics Speech Signal Process. ASSP-26 (4) (1978) 334–338.